DOCUMENT RESUME

ED 420 716 TM 028 438

AUTHOR Fouladi, Rachel T.

TITLE Type I Error Control of Two-Group Multivariate Tests on Means under Conditions of Heterogeneous Correlation Structure.

PUB DATE 1998-04-00

NOTE 28p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Diego, CA, April 13-17, 1998).

PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.

DESCRIPTORS *Correlation; Monte Carlo Methods; *Multivariate Analysis; Robustness (Statistics); Simulation; Tables (Data)

IDENTIFIERS Hotellings t; Mean (Statistics); *Type I Errors

ABSTRACT

A variety of approaches have been suggested by which to assess the equality of population mean vectors under conditions of population covariance matrix homogeneity and heterogeneity. The nonrobustness of commonly used multivariate tests of means to population covariance matrix heterogeneity has been long documented. However, most studies have examined the performance characteristics of the statistical procedures under conditions of heterogeneous covariance structure by simulating heterogeneity in the structure of the variances. The only study that examined performance under heterogeneous covariance structure by simulating heterogeneity in the correlations concluded that there was little difference in the performance characteristics of standard multivariate means under conditions of variance homogeneity and correlation heterogeneity (T. Beasley and J. Sheehan, 1994); this study, however, only examined the performance of the procedures under equal sample sizes. This paper assesses the Type I error control of standard and alternative multivariate tests of means under homogeneous and heterogeneous correlation structure for a full range of sample size conditions. This paper focuses on the performance of multivariate tests on means in the two-group case. A Monte Carlo simulation experiment was conducted. Findings show that the "F" based on Hotelling's T-squared is robust to between groups differences in correlation matrices under equal and unequal sample size conditions as long as the difference in the magnitude of the correlations is not extremely large, no matter what the sample size conditions or the number of variables under study. Differences between the performance profiles of the standard multivariate means test procedure and available alternative procedures are discussed. An appendix provides a chart of observed percent bias on Type I error control of multivariate tests on means. (Contains 5 tables and 22 references.) (Author/SLD)

Type I error control of two-group multivariate tests on means under conditions of heterogeneous correlation structure

Rachel T. Fouladi

Department of Educational Psychology
University of Texas at Austin
Austin, Texas 78712-1296

Abstract. Over the decades, a variety of approaches have been suggested by which to assess the equality of population mean vectors under the condition of population covariance matrix homogeneity and heterogeneity. The nonrobustness of commonly used multivariate tests of means to population covariance matrix heterogeneity has been long documented. However, most studies have examined the performance characteristics of the statistical procedures under conditions of heterogeneous covariance structure by simulating heterogeneity in the structure of the variances. The only study which examined performance under heterogeneous covariance structure by simulating heterogeneity in the correlations concluded that there was little difference in the performance characteristics of standard multivariate tests of means under conditions of variance homogeneity and correlation heterogeneity (Beasley & Sheehan, 1994); this study, however, only examined the performance of the procedures under equal sample sizes. The present paper assesses the Type I error control of standard and alternative multivariate tests of means under homogeneous and heterogenous correlation structure for a full range of sample size conditions. This paper focuses on the performance of multivariate tests on means in the two group case.

Subject descriptors: Hotelling's T-squared, multivariate tests on means, heterogeneity of covariance matrices, heterogeneity of correlation matrices, Type I error, MANOVA.

Introduction

Multivariate data analytic procedures are widely used by researchers in many different disciplines. Multivariate questions that are of interest to many researchers include comparisons of mean, correlation, and covariance structure between experimental and intact groups. Even though there are a wide variety of available data analytic techniques, the procedures which are most commonly used to compare the mean structure of several groups on several variables include parametric multivariate analysis of variance (MANOVA) and related discriminant function analysis techniques, where assumptions are that the observations in the comparison groups are obtained from multivariate normal populations with homogeneous covariance matrices.

As is well known, not all statistical assumptions are realistic nor are all procedures robust to assumption violations. Thus, the question of the tenability of assumptions is an important point of consideration. While the question of whether data can be assumed to be obtained from multivariate normal populations has received much attention in recent years (e.g., Micceri, 1989), the question of whether data can be assumed to be obtained from populations with homogenous covariance structures has received far less attention.

The question of the tenability of the assumption of homogenous covariance matrices is especially salient when testing differences between intact groups, but can also be an issue with experimental groups, with the heterogeneity of covariance matrices manifesting in two basic ways, (a) the variances of some or all of the variables are different, and/or (b) some or all of the variables are correlated differently in at least two of the groups under study. Though the tenability of the assumption of homogeneity of covariance matrices is not generally discussed, some conditions under which commonly used multivariate tests of means are nonrobust to population covariance matrix heterogeneity have been documented (e.g., Hakstian, Roed, Linn, 1979; Holloway and Dunn, 1967; Olson, 1974).

Importantly however, with the exception of Beasley and Sheehan (1994), most studies examining the performance characteristics of the multivariate tests on means under conditions of heterogeneous covariance structure simulate heterogeneity in the structure of the variances, not heterogeneity in the structure of the correlations. As such, the impact of heterogeneous patterns in the variable variances has been widely studied (Algina

Presented at the American Educational Research Association Annual Meeting, April 11-17, 1998, San Diego, CA

& Oshima, 1990; Algina, Oshima, & Tang, 1991; Algina & Tang, 1988; Everitt, 1979; Hakstian, Roed, & Lind, 1979; Holloway & Dunn, 1967; Hopkins & Clay, 1963; Kim, 1992; Mardia, 1971; Subrahmaniam & Subrahmaniam, 1973; Yao, 1965); in contrast, the impact of heterogeneous correlation patterns has been studied very little.

Over the decades, a variety of approaches have been suggested by which to assess the equality of population mean vectors under the condition of population covariance matrix homogeneity and heterogeneity. Test statistics which have been proposed for use under conditions of heterogeneity of covariance matrices include procedures suggested by James (1954), Yao (1965), Johansen (1980), Nel and van der Merwe (1986), and Kim (1992). Studies examining the performance of these alternative techniques have established the improved performance of these procedures over the standard parametric procedures under conditions of heterogeneous patterns in the variable variances; no study has examined the relative performance of these techniques under heterogeneous correlation patterns.

The present study assesses the Type I error control of standard and alternative multivariate tests of means under homogenous and heterogenous correlation structure for a full range of sample size conditions. This paper focuses on the performance of multivariate tests on means in the two group case. The procedures under study and reviewed in the following section include the standard parametric F statistic based on Hotelling's T-squared, and alternative test statistics suggested by James (1954), Yao (1965), Johansen (1980), Nel and van der Merwe (1986), and Kim (1992), and recommended for use under conditions of heterogeneous covariance matrices.

Test procedures examined

Consider n_i independent identically distributed observation vectors $\mathbf{x}_1,\ldots,\mathbf{x}_{n_i}$ obtained from a p-dimensional multivariate normal population, with $p \times 1$ population mean vector $\boldsymbol{\mu}_i$, non-singular $p \times p$ population covariance matrix $\boldsymbol{\Sigma}_i$ and population correlation matrix $\mathbf{P}_i$. Let $\overline{\mathbf{x}}_i$ and $\mathbf{S}_i$ represent the sample mean and the sample covariance matrix for the ith group (i=1,2) and $\mathbf{S} = (n_1 + n_2 - 2)^{-1}[(n_1 - 1)\mathbf{S_1} + (n_2 - 1)\mathbf{S_2}]$.

Many multivariate tests of the null hypothesis on the equality of two population mean vectors are formulated either as functions of the scalar quantities $T_E^2 = (\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2)'\left[n_1^{-1}\mathbf{S} + n_2^{-1}\mathbf{S}\right]^{-1}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2)$ or $T_U^2 = (\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2)'\left[n_1^{-1}\mathbf{S}_1 + n_2^{-1}\mathbf{S}_2\right]^{-1}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2)$, where statistics which are functions of T_E^2 are typically based on the assumption of the equality of the two population covariance matrices, whereas statistics which are function of T_U^2 are not typically based on the assumption of covariance matrix homogeneity.

For a test of the null hypothesis on the equality of two population mean vectors, the standard parametric statistic (Hotelling, 1951), based on the assumption that observations are independent identically distributed from multivariate normal populations with homogeneous population covariance matrices, is $F_H = [(n_1 + n_2 - 2)p]^{-1}(n_1 + n_2 - p - 1)T_E^2$. This statistic has an F distribution with p and $n_1 + n_2 - p - 1$ degrees of freedom.

A wide variety of alternative test statistics have been proposed. The alternative approaches under consideration in this paper are procedures suggested by James (1954), Yao (1965), Johansen (1980), Nel and van der Merwe (1986), and Kim (1992). These procedures all assume that observations are independent identically distributed from multivariate normal populations; however they do not assume that the populations have homogeneous covariance matrices.

For a test of the null hypothesis that two population mean vectors are equal, James (1954) expressed the critical value for T_U^2 as a series of terms in descending order of magnitude. The 1st order approximation of the critical value is given by $c(A + cB)$ where c is the $1-\alpha$ percentile point of the central chi-square distribution with p degrees of freedom, $\mathbf{A}_i = (n_i)^{-1}\mathbf{S}_i$, $\mathbf{V} = \sum_{i=1}^{2}\mathbf{A}_i$, $A = 1 + (2p)^{-1}\sum_{i=1}^{2}(n_i - 1)^{-1}tr^2(\mathbf{V}^{-1}\mathbf{A}_i)$, and

$B = [p(p+2)]^{-1}\sum_{i=1}^{2}(n_i - 1)^{-1}[tr(\mathbf{V}^{-1}\mathbf{A}_i)^2 + .5tr^2(\mathbf{V}^{-1}\mathbf{A}_i)]$. James' 2nd order approximation to the critical value is given by the sum of James' 1st order critical value and

$$
\begin{aligned}
2h_2(a) = {} & \frac{1}{16}\left(1-\frac{p-2}{c}\right)\left[\sum \frac{1}{\nu_i}\left(2\chi_4[i|i]+(\chi_4+\chi_2)[i]^2\right)\right]^2 \\
& -\sum \nu_i^{-2}\left((2\chi_4+\chi_2)[i|i]+\chi_4[i]^2\right) \\
& +\sum \nu_i^{-2}\left(2(3\chi_4+\chi_2)[i|i|i]+(5\chi_4+\chi_2)[i|i][i]+(\chi_4+\chi_2)[i]^3\right) \\
& -\sum (\nu_i\nu_j)^{-1}\left(2\chi_4[i|i|j|j]+(2\chi_4+\chi_2)[i|j|i|j]+(3\chi_4+\chi_2)[i|i|j][j]+\chi_4[i|j]^2+(\chi_4+\chi_2)[i|j][i][j]\right) \\
& +(\chi_2-1)\sum \nu_i^{-2}\left(2\chi_4[i|i|i]+(\chi_4+\chi_2)[i|i][i]\right) \\
& -(\chi_2-1)\sum (\nu_i\nu_j)^{-1}\left(2\chi_4[i|i|j|j]+(\chi_4+\chi_2)[i|i|j][j]\right) \\
& -\frac{1}{8}\sum (\nu_i\nu_j)^{-1}\left(2(\chi_4-\chi_2)[i|i]+(\chi_4-1)[i]^2\right)\left(2\chi_4[j|j]+(\chi_4+\chi_2)[j]^2\right) \\
& -\frac{1}{3}\sum \nu_i^{-2}\left(2(4\chi_6+\chi_4+\chi_2)[i|i|i]+3(2\chi_6+\chi_4)[i|i][i]+(\chi_6+\chi_4+\chi_2)[i]^3\right) \\
& +\frac{1}{16}\sum (\nu_i\nu_j)^{-1}\left(\begin{aligned} & 32\chi_8[i|i|j|j]+8(2\chi_8+2\chi_6+\chi_4+\chi_2)[i|j|i|j]+16(2\chi_8+\chi_6+\chi_4)[i|i|j][j] \\ & +4(\chi_8-\chi_6)[i|i][j|j]+8(\chi_8+\chi_6)[i|j]^2+4(\chi_8-\chi_4)[i|i][j]^2 \\ & +8(\chi_8+\chi_6+\chi_4+\chi_2)[i|j][i][j]+(\chi_8+\chi_6-\chi_4-\chi_2)[i]^2[j]^2 \end{aligned}\right),
\end{aligned}
$$

where c is the $1-\alpha$ percentile point of the central chi-square distribution with p degrees of freedom, and $\chi_2 = c(p)^{-1}$, $\chi_{2s} = \chi_{2(s-1)}[p+2(s-1)]^{-1}$ for $s>1$, $\nu_i = n_i - 1$, $[i] = tr(\mathbf{V}^{-1}\mathbf{A}_i)$, $[i,j] = tr(\mathbf{V}^{-1}\mathbf{A}_i\mathbf{V}^{-1}\mathbf{A}_j)$, $[i,j,k] = tr(\mathbf{V}^{-1}\mathbf{A}_i\mathbf{V}^{-1}\mathbf{A}_j\mathbf{V}^{-1}\mathbf{A}_k)$, and $[i,j,k,l] = tr(\mathbf{V}^{-1}\mathbf{A}_i\mathbf{V}^{-1}\mathbf{A}_j\mathbf{V}^{-1}\mathbf{A}_k\mathbf{V}^{-1}\mathbf{A}_l)$.

Yao (1965) suggested a test statistic based on a transformation of T_U^2. This test statistic is $F_Y = (pf_2)^{-1}(f_2 - p + 1)T_U^2$ where $f_2^{-1} = \sum_{i=1}^{2}(n_i-1)^{-1}\left(w_i/T_U^2\right)^2$, $w_i = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'\mathbf{V}^{-1}\mathbf{A}_i\mathbf{V}^{-1}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)$. For Yao's F_Y, critical values are obtained from the *F* distribution with p and $f_2 - p + 1$ degrees of freedom.

Johansen (1980) proposed the test statistic $F_{Jo} = c_1^{-1}T_U^2$ where $c_1 = p + 2C - 6C(p+1)^{-1}$ and $C = 5\sum_{i=1}^{2}(n_i-1)^{-1}[tr(\mathbf{V}^{-1}\mathbf{A}_i)^2 + tr^2(\mathbf{V}^{-1}\mathbf{A}_i)]$. For a test of the null hypothesis, the reference distribution of Johanson's F_{Jo} is the *F* distribution with p and $p(p+2)/3A$ degrees of freedom.

The Nel and van der Merwe (1986) test statistic is $F_N = (pf_3)^{-1}(f_3 - p + 1)T_U^2$ where $f_3 = (tr\mathbf{V}^2 + tr^2\mathbf{V})\sum_{i=1}^{2}(n_i-1)^{-1}(tr\mathbf{A}_i^2 + tr^2\mathbf{A}_i)$. Nel and van der Merwe's F_N is referred to the *F* distribution with p and $f_3 - p + 1$ degrees of freedom.

Kim (1992) suggested an alternative test statistic. This statistic is

$F_K = (c_2 m f_2)^{-1}(f_2 - p + 1)(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'\mathbf{A}^{-1}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)$, where

$\mathbf{A} = \mathbf{A}_1 + r^2\mathbf{A}_2 + 2r\mathbf{A}_2^{1/2}(\mathbf{A}_2^{-1/2}\mathbf{A}_1\mathbf{A}_2^{-1/2})^{1/2}\mathbf{A}_2^{1/2}$, $r = \left|\mathbf{A}_1\mathbf{A}_2^{-1}\right|^{1/(2p)}$, $c_2 = \sum_{j=1}^{p} L_j^2 \Big/ \sum_{j=1}^{p} L_j$,

$m=\left(\sum_{j=1}^{p} L_j\right)^2 \bigg/ \sum_{j=1}^{p} L_j^2$, $L_j=(d_j+1)\big/(d_j^{1/2}+r)^2$, and d_j is the jth eigenvalue of $\mathbf{A}_1\mathbf{A}_2^{-1}$. Kim's F_K is referred to the F distribution with p and f_2-p+1 degrees of freedom.

Relevant Monte Carlo research

Numerous researchers (e.g., Algina & Oshima, 1990; Algina, Oshima, & Tang, 1991; Algina & Tang, 1988; Everitt, 1979; Hakstian, Roed, & Lind, 1979; Holloway & Dunn, 1967; Hopkins & Clay, 1963; Kim, 1992; Mardia, 1971; Subrahmaniam & Subrahmaniam, 1973; Yao, 1965) have examined the performance of test statistics which enable the comparison of the mean vectors of two populations. Even though many studies have examined the performance of two-group multivariate tests on means under conditions of heterogeneous population covariance matrices, the studies examining the performance of the test statistics under heterogeneity of covariance matrices have simulated heterogeneity of the covariance matrices only by varying between groups the population variances of the underlying variables, not by varying between groups the population correlation structure of the underlying variables.

In general the research on the two group multivariate test statistics on means have shown that the standard F based on Hotelling's T-squared is relatively robust under equal sample size conditions, however under unequal sample size conditions if the group with the smaller sample size has the smaller variances then the standard F procedure is conservative, and if the group with the smaller sample size has the larger variances then the standard F is liberal. Furthermore, the alternative procedures provide improved Type I error control in general (c.f., Coombs, Algina, and Oltman, 1996).

Beasley and Sheehan (1994) conducted a study on the impact of homogeneous variances and heterogeneous covariances on standard MANOVA procedures, of which the standard parametric two-group Hotelling T-squared procedure is a special case. For the conditions they examined, they determined that when the variances were equal the presence of unequal covariances did not impact the performance of the MANOVA procedures, that is, they found that the presence of heterogeneous correlation matrices did not impact the performance of the MANOVA procedures. Their study, however, only examined the performance of the MANOVA procedure under conditions of equal sample sizes.

From the results of studies simulating only heterogeneity of variances, that MANOVA procedures are fairly robust to moderate heterogeneity of variances under equal sample sizes is well known; however, it is also well known that robustness does not obtain under unequal sample sizes. Thus, even though Beasley and Sheehan's results indicate that MANOVA procedures are robust to heterogeneity of correlation structure under equal sample sizes, one might expect that when sample sizes are unequal, the MANOVA procedures will not perform well for comparisons of mean vectors between groups from populations with heterogeneous correlation matrices.

Methods

A Monte Carlo simulation experiment was conducted in order to compare the Type I error control of procedures available to assess whether the mean vectors for two groups are different in the population under conditions of heterogeneous correlation structure. The relative performance of the standard Hotelling T-squared F statistic and alternative James 1st order, James 2nd order, Yao, Johansen, Nel, and Kim procedures were assessed.

A stand-alone FORTRAN computer program, implementing the standard parametric and alternative tests, was written for this study. This program was written by the author; some IMSL routines were used.

Design

Data are generated from multivariate normal populations where all the variables have mean 0 and unit variance. The Type I error control of the standard and alternative statistics are examined under the following conditions.

Population correlation structure, $\mathbf{P}_1$ and $\mathbf{P}_2$: Data are from populations where variables are homogeneously intercorrelated. Data in group 1 are always generated from populations where variables are uncorrelated. Data in group 2 are generated from populations where the magnitude of the population correlations are .0, .1, .3, .5, .7, or .9.

Number of variables, p: Data are from p-variate multinormal populations, where p equals 4, 8, or 12.

Sample size, n_1 and $n_1{:}n_2$: Data for group 1 are generated at specific ratios of sample size to number of variables; sample size for group 1, n_1, include the conditions (p+1), $2p$, $4p$, $10p$, $20p$, $40p$. Data for group 2 are generated as a function of the sample size in group 1. The ratios of sample size for group 1 to sample size for group 2, n_1:n_2, are: 1:1, 1:2, 1:4, 2:1, 4:1.

In typical simulations where observations are simulated to be independent identically distributed, a sequence of independent uniform variates (usually real-valued between 0 and 1) are first generated and then transformed in an appropriate way. The method of Kinderman and Ramage (1976) was used to generate data from multivariate normal distributions with the specified correlation structure.

Hypotheses were tested at one level of nominal Type I error: α =.05. For each data set, the test statistics and critical values necessary for assessing the equality of mean vectors were calculated; the decisions for the procedures were recorded.

The four factors $\mathbf{P}_2$, p, n_1 and n_1:n_2 are fully crossed, resulting in a 6x3x6x5 factorial design. Each condition is replicated 10,000 times.

Measures of performance

Under each condition, the rejection frequency for each statistic is observed. For each condition, the number of rejections obtained for each test is tabulated and transformed into proportion rejected. Under each condition, the empirical rejection rate, $\alpha_{\text{Empirical}}$, for each statistic is observed. For each cell, the bias and percent bias results are obtained. For each cell, the percent bias ($B_{\%}$) of the observed empirical rejection rate from the expected rejection rate, α_{Nominal}, is obtained where $B\% = 100(\alpha_{\text{Empirical}} - \alpha_{\text{Nominal}}) / \alpha_{\text{Nominal}}$. Factorial analysis of variance designs are used to determine the influence of the different factors on the pattern of decisions. Chi-squared goodness of fit values based on a normal approximation to the binomial are also computed; from this information, whether a procedure controls Type I error at the nominal level is assessed. Percent bias are also examined using the Bradley (1978) and Robey, and Barcikowski (1992) guidelines for what constitutes acceptable departures of empirical rejection rates from the nominal rejection rates.

Results

Empirical Type I error rate performance of each test statistic was assessed with 10,000 replications under every cell in the design. Bradley (1978) asserted that many researchers are unreasonably generous when defining acceptable departures of empirical alpha from the nominal level. He held that the departure of empirical alpha from the nominal level was "negligible" if empirical alpha was within $\alpha \pm \frac{1}{10}\alpha$ according to a 'fairly stringent criterion', and $\alpha \pm \frac{1}{2}\alpha$ according to the "most liberal criterion that [he] was able to take seriously" which in the remainder of his article he referred to as the 'liberal criterion'. Robey and Barcikowski (1992) supplement the guidelines provided by Bradley for defining acceptable departures from the nominal level, providing an 'intermediate criterion' of $\alpha \pm \frac{1}{4}\alpha$, and a 'very liberal criterion' of $\alpha \pm \frac{3}{4}\alpha$.

Appendix A details the percent bias ($B_{\%}$) results of each of the procedures. Inspection of the results showed that no procedure consistently controls empirical Type I error rates within any of the Bradley, Robey and Barcikowski (BRB) criteria, and that patterns vary across levels of $\mathbf{P}_2$, p, n_1 and n_1:n_2.

Assessing whether there is a significant difference between the Type I error control of the procedures under study and the pattern of influence of P2, p, n1 and n1:n2.

A factorial multivariate analysis of variance was conducted on the departures of empirical rejection rates from the nominal level; procedure type was parameterized as a repeated measures factor and $\mathbf{P}_2$, p, n_1 and n_1:n_2 were between subjects factors. The multivariate test for procedure type yielded p<.001. All multivariate tests of interaction effects involving procedure type included in the model yielded p<.001; the five way interaction effect was not tested as there was only one summary empirical rejection rate per cell. Follow-up factorial analyses were conducted for each test procedure. All tests of main, two-way and three-way interaction effects yielded p<.001, with the exception of $p \times n_1$ for *Hotelling's F* (p=.031), $p \times \mathbf{P}_2$ for *Kim* (p=.019), $p \times n_1 \times n_1{:}n_2$ for *Hotelling's F* (p>.05), $p \times n_1 \times \mathbf{P}_2$ for *Kim* (p>.05), $n_1 \times n_1{:}n_2 \times \mathbf{P}_2$ for *James1*, *James2*, *Yao*, *Johanson*, and *Kim* (p>.05). As such there is a significant difference in the Type I error control of the different statistical procedures, and this control varies across levels of $\mathbf{P}_2$, p, n_1 and n_1:n_2.

Assessing the departure of the empirical Type I error rate from the nominal level

The Type I error control of the procedures was analyzed overall and across levels of $\mathbf{P}_2$, p, n_1 and n_1:n_2. Chi-square results and summary statistics on the percent bias of the different procedures are shown in Tables 1-5.

Chi-square goodness of fit

Chi-square goodness of fit values to assess the departure of the empirical Type I error rate from the nominal level were computed for every factorial cell for each test statistic. Chi-squares were summed to yield composite chi-square goodness of fit tests.

Chi-square results show that overall none of the procedures controlled empirical Type I error rates at the nominal level across all the conditions examined. Analyses were also conducted to determine whether the statistical significance of the departures of the empirical Type I error rate from the nominal level varied at different levels of $\mathbf{P}_2$, p, n_1 and n_1:n_2.

For the types of correlation pattern $\mathbf{P}_2$ examined, the chi-square results show that the empirical rejection rates of the standard Hotelling *T*-squared *F* statistic were not significantly different from the nominal level when $\mathbf{P}_1$ equals $\mathbf{P}_2$, whereas, the empirical rejection rates of the alternative procedures were significantly different from the nominal level under this condition. For $\mathbf{P}_1$ equal to $\mathbf{P}_2$, the magnitude of the chi-square results show that *Hotelling's F* evidenced the best overall control of the empirical rejection rate within the nominal level, followed by *Nel, James2, Johanson, Kim, Yao,* and *James2*. For all other conditions where $\mathbf{P}_1$ did not equal $\mathbf{P}_2$, the empirical rejection rates of the standard and alternative procedures were significantly different the nominal level. For $\mathbf{P}_1$ unequal to $\mathbf{P}_2$, when the magnitude of the correlations in group 2 were all .1, *Hotelling's F* evidenced the best overall control of the empirical rejection rate within the nominal level, followed by *Nel, James2, Johanson, Kim, Yao,* and *James2*. However when the magnitude of the correlations in group2 were all .3, the order from least to greatest overall departure from the nominal level was *Nel, James2, Johanson, Hotelling's F, Kim, Yao,* and *James1*. For correlations of .5 in group 2, *Nel* showed the least overall departure from the nominal level, followed by *James2, Kim, Johanson, Yao, James1,* and *Hotelling's F;* for correlations of .7, *Nel* was followed by *Kim, James2, Johanson, Yao, James1,* and *Hotelling's F*; and for correlations of .9, *Kim* was followed by *Nel, James2, Yao, Johanson, James1,* and *Hotelling's F.*

Under the levels of p examined, the chi-square results indicate the empirical rejection rates for all the procedures were significantly different from the nominal level under every level of p. Though significantly different from the nominal level, at p equal 4, the magnitude of the chi-square results suggest *Nel* showed the least overall departure from the nominal level, followed by *James2, Johanson, Kim, Yao, James1*, and then the standard *Hotelling F.* At p equal 8, *Nel* still showed the least overall departure from the nominal level, followed by *James2, Kim, Johanson, Yao, James1*, and then the standard *Hotelling F.* At p equal 12, *Nel* continued to show the least overall departure from the nominal level, followed by *Kim, James2, Johanson, Yao, James1*, and then the standard *Hotelling F.*

The chi-square results indicate the empirical rejection rates of all the procedures were significantly different from the nominal level for n_1 equal to $(p+1)$, $2p$, and $4p$; however for n_1 equal to $10p$, $20p$, and $40p$, some procedures had empirical rejection rates that were not significantly different from the nominal level. Though significantly different from the nominal level, at n_1 equal to $p+1$, the chi-square results indicate *Nel* showed the least overall departure from the nominal level followed by *Kim, James2, Johanson, Yao, Hotelling's F,* and *James2*; at n_1 equal to $2p$, the order from least to greatest overall departure from the nominal level was *Kim, James2, Nel, Johanson, Yao, James2*, followed by *Hotelling's F*; and at n_1 equal to $4p$, James2 showed the least overall departure from the nominal level followed by *Johanson, Yao, James2, Kim, Nel,* and *Hotelling's F.* At of n_1 equal to $10p$, *James2, Johanson,* and *Yao* had empirical rejection rates that were not significantly different from the nominal level. At of n_1 equal to $20p$, *James1, James2, Yao,* and *Johanson* had empirical rejection rates that were not significantly different from the nominal level. At of n_1 equal to $40p$, *James1, James2, Yao, Johanson,* and *Nel* had empirical rejection rates that were not significantly different from the nominal level.

Under the levels of n_1:n_2 examined, the chi-square results indicate empirical rejection rates were significantly different from the nominal level under every level of n_1:n_2 for every statistic. Though with an empirical rejection rate significantly different from the nominal level, at n_1:n_2 equal to 1, *Nel* showed the least overall departure from the nominal level, followed by *Kim, Yao, Hotelling's F, James2, Johanson,* and *James1*. At n_1:n_2 equal to 1:2, *Kim* showed the least overall departure from the nominal level, followed by *Nel, James2, Yao, Johanson, James1,* and *Hotelling's F.* At n_1:n_2 equal to 1:4, *Nel* showed the least overall departure from the nominal level, followed by *James2, Kim, Johanson, Yao, James1,* and *Hotelling's F.* At n_1:n_2 equal to 2:1, *James2* showed the least overall departure from the nominal level, followed by *Yao, Kim, Johanson, Nel, Hotelling's F,* and *James1*. At $n_1{:}n_2$ equal to 4:1, *James2* showed the least overall departure from the nominal level, followed by *Kim,, Johanson, James1, Yao, Hotelling's F,* and *Nel.*

Bradley, Robey, and Barcikowski guidelines and percent bias

According to the Bradley, Robey, and Barcikowski (BRB) guidelines for what constitutes acceptable levels of departure of empirical Type I error rates from the nominal level, procedures which control empirical rejection rates within $\alpha \pm \frac{1}{10}\alpha$ are described as providing "stringent" Type I error control, within $\alpha \pm \frac{1}{4}\alpha$ as providing "intermediate" control, within $\alpha \pm \frac{1}{2}\alpha$ as providing "liberal" control, and within $\alpha \pm \frac{3}{4}\alpha$ as providing "very liberal" control. Judgments are based on whether procedures *consistently* provided control of empirical rejection rates across the conditions, i.e., whether they provided control within the level specified across every cell under consideration; as such, judgments are based on whether the minimum and maximum percent bias of a given procedure is within the BRB guidelines across the conditions under consideration.

As indicated earlier, no procedure consistently satisfies the BRB critieria for acceptable Type I error control across all conditions, nor does any procedure consistently control empirical rejection rates within $\alpha \pm \alpha$.

For the different types of correlation pattern $\mathbf{P}_2$ examined, the F statistic based on the standard Hotelling T-squared F statistic consistently provided stringent Type I error control for the conditions where $\mathbf{P}_1$ equals $\mathbf{P}_2$; *Nel* controlled empirical rejection rates within $\alpha \pm \alpha$ under this condition. For conditions where $\mathbf{P}_1$ did not equal $\mathbf{P}_2$, with the magnitude of the population correlation coefficients in group 2 equal to .1, the standard Hotelling F statistic provided consistent control of the empirical rejection rate within the liberal criterion; *Nel* controlled empirical rejection rates within $\alpha \pm \alpha$ under this condition. For the conditions where the magnitude of the population correlation coefficients in group 2 are all .3, .5, or .7, *Nel* consistently controlled empirical rejection rates within $\alpha \pm \alpha$. For the conditions where the magnitude of the population correlation coefficients in group 2 are .9, no procedure consistently controlled the empirical rejection rate within the BRB criteria for acceptable Type I error control or within $\alpha \pm \alpha$.

At p equal to 4, no procedure consistently controlled empirical rejection rates within the BRB criteria; however, *Nel* did consistently control empirical rejection rates within $\alpha \pm \alpha$. At p equal to 8 or 12, no procedure controlled empirical rejection rates within the BRB critieria or $\alpha \pm \alpha$.

Under the levels of n_1 examined, no procedure consistently controls the empirical rejection rates within the BRB critieria for acceptable departures of empirical rejection rates from the nominal level or within $\alpha \pm \alpha$ for n_1 equal to $(p+1)$. For n_1 equal to $2p$, though *James2* and *Kim* consistently control the empirical Type I error rate within $\alpha \pm \alpha$ which none of the other procedures do. For n_1 equal to $4p$, *James2* and *Johanson* consistently control the empirical rejection rate within the intermediate criterion, *James1* and *Yao* control the empirical rejection rate within the liberal criterion, *Kim* provides control within the very liberal criterion, and *Nel* within $\alpha \pm \alpha$. For n_1 equal to $10p$, *James1, James2, Yao,* and *Johanson* control the empirical rejection rate within the intermediate criterion, and *Nel* and *Kim* provide control within the liberal criterion. For n_1 equal to $20p$, *James1, James2, Yao, Johanson, Nel* and *Kim* control the empirical rejection rate within the intermediate criterion. For n_1 equal to $40p$, *James1, James2, Yao,* and *Johanson* control the empirical rejection rate within the stringent criterion, and *Nel* and *Kim* provide control within the intermediate criterion.

Under the levels of $n_1{:}n_2$ examined, the only procedure which consistently controls the empirical rejection rates within any of the BRB critieria for acceptable departures of empirical rejection rates from the nominal level for $n_1{:}n_2$ equal to 1:1 is *Kim*, though the *Nel* procedure does provide control within $\alpha \pm \alpha$. No procedure controls empirical Type I error rates within any of the BRB criteria for $n_1{:}n_2$ equal to 1:2 or 1:4. However for $n_1{:}n_2$ equal to 2:1, *James2* controls empirical rejection rates within the very liberal criterion, and the standard *Hotelling F, Yao,* and *Nel* control the empirical rejection rate within $\alpha \pm \alpha$.

Conclusions

The findings in this paper on the two group multivariate tests on means show that the F based on Hotelling's T-squared is robust to between groups differences in correlation matrices under equal and unequal sample size conditions as long as the difference in the magnitude of the correlations is not extremely large, no matter what the sample size conditions or the number of variables under study. However under moderate to large between group differences in the correlation structure of the variables under study, the standard Hotelling T-squared F

procedure does not yield empirical Type I error rates that are consistently close to the nominal level. Under moderate to large differences in the population correlation matrices, if sample sizes are equal, the magnitude of the differences is not extreme, and the number of variables under study is not particularly large, the standard *F* procedure performs quite well; however if sample sizes are unequal, no matter what the sample size or the sample size to number of variables ratio, if the group with the smaller sample size has the variables which are more strongly intercorrelated then the standard *F* procedure is conservative, and if the group with the smaller sample size has the variables which are more weakly intercorrelated then the standard *F* is liberal.

The results of the current study on the performance of the Hotelling *T*-square *F* statistic are consistent with the literature on the heterogeneity of covariance matrices which simulated only heterogeneity of variances and not heterogeneity of the correlations. The literature on the impact of heterogeneity of variances indicates that when the larger groups have the variables with the larger variances and the smaller groups have the variables with the smaller variances, then the standard parametric procedures for multivariate tests on means are conservative; the literature also indicates that that when the larger groups have the variables with the smaller variances and the smaller groups have the variables with the larger variances, then the standard procedures for multivariate tests on means are liberal. Thus, when there is a positive relationship between sample size and the generalized variance of the groups, the standard procedures are conservative; and when there is a negative relationship, the procedures are liberal. For the conditions simulated in the present study, the generalized variance of a group with variables that are strongly intercorrelated is smaller than the generalized variance of a group with variables that are weakly intercorrelated. Thus, just as when only variance heterogeneity is simulated, when correlation heterogeneity is simulated, if there is a positive relationship between sample size and the generalized variance of the groups then the standard parametric multivariate means test procedure is conservative; and if there is a negative relationship then the procedure is liberal.

Results showed clear differences between the performance profiles of the standard multivariate means test procedure and available alternative procedures. Unlike the Hotelling's *T*-squared *F* procedure, the alternative procedures showed extremely good Type I error control at moderate to large sample sizes no matter what the number of variables under study, the magnitude of the between group differences in the correlation matrices, or the relationship between sample size and the generalized variance. Differences between the alternative procedures were mainly in terms of sample size requirements to yield acceptable Type I error control, with *James2* and *Johanson* showing the fastest convergence to acceptable Type I error rates, followed by *James1*, *Yao*, *Nel* and *Kim*.. Importantly for researchers analyzing data from small sample research, none of the alternative procedures had acceptable Type I error rates for the smallest level of sample size; under these conditions a researcher is well advised to have equal sample sizes and use the standard parametric techniques. However for the analysis of data sets where sample sizes are unequal and the sample size to number of variables ratio is not extremely small, alternative techniques are preferred.

Author Note

Rachel T. Fouladi (Ph.D., University of British Columbia, 1996) is Assistant Professor in Research Methodology and Data Analysis at the Department of Educational Psychology at the University of Texas at Austin. Correspondence concerning this article should be addressed to the author at: University of Texas at Austin, Dept of Educational Psychology, SZB 504, Austin, TX 78712-1296 U.S.A. Phone 512-471-4155, Fax 512-471-1288, Email rachel.fouladi@mail.utexas.edu, http://www.edb.utexas.edu/faculty/fouladi/.

References

Algina, J., & Oshima, T. C. (1990). Robustness of the independent samples Hotelling's T^2 to variance-covariance heteroscedasticity when sample sizes are unequal and in small ratios. Psychological Bulletin, 108, 308-313.

Algina J, Oshima, T. C., & Tang, K. L. (1991). Robustness of Yao's, James', and Johansen's tests under variance-covariance heteroscedasticity and nonnormality. Journal of Educational Statistics, 16, 125-139.

Algina, J., & Tang, K. L. (1988). Type I error rates for Yao's and James' tests of equality of mean vectors under variance-covariance heteroscedasticity. Journal of Educational Statistics, 13, 281-290.

Beasley, T.M., & Sheehan, J.K. (1994). Choosing a MANOVA test statistic when covariances are unequal. Paper presented at the Annual Meeting of the Midwestern Educational Research Association, Chicago, IL, October 1994.

Bradley, J.V. (1978). Robustness? British Journal of Mathematical and Statistical Psychology, 34, 144-152.

Coombs, W. T., Algina, J., & Oltman, D. O. (1996). Univariate and multivariate omnibus hypothesis tests selected to control type 1 error rates when population variances are not necessarily equal. Review of Educational Research, 66, 137-179.

Everitt, B. S. (1979). A Monte Carlo investigation of the robustness of Hotelling's one- and two-sample T^2 tests. Journal of the American Statistical Association, 74, 48-51.

Hakstian, A. R., Roed, J. C., & Lind, J. C. (1979). Two-sample T^2 procedure and the assumption of homogeneous covariance matrices. Psychological Bulletin, 86, 1255-1263.

Holloway, L. N., & Dunn, O. J. (1967). The robustness of Hotelling's T^2. Journal of the American Statistical Association, 62, 124-136.

Hopkins, J. W., & Clay, P. P. F. (1963). Some empirical distributions of bivariate T^2 and homoscedasticity criterion M under unequal variance and leptokurtosis. Journal of the American Statistical Association, 58, 1048-1053.

Hotelling, H. (1951). A generalized T test and measure of multivariate dispersion. In J. Neyman (Ed.), Proceedings of the second Berkeley Symposium on Mathematical Statistics and Probability (pp. 23-41). Berkeley: University of California Press.

James, G. S. (1954). Tests of liner hypothesis in univariate and multivariate analysis when the ratios of population variances are unknown. Biometrika, 41, 19-43.

Johansen, S. (1980). The Welch-James approximation to the distribution of the residual sum of squares in a weighted linear regression. Biometrika, 67, 85-92.

Kim, S.-J. (1992). A practical solution to the multivariate Behrens-Fisher problem. Biometrika, 79, 171-176.

Kinderman, A.J., & Ramage, J.G. (1976). Computer generation of normal random variables. Journal of the American Statistical Association, 71, 893-896.

Mardia, K. V. (1971). The effect of nonnormality on some multivariate tests and robustness to nonnormality in the linear model. Biometrika, 58, 105-121.

Nel, D. G., and van der Merwe, C. A. (1986). A solution to the multivariate Behrens-Fisher problem. Communications in Statistics-Theory and Methods, 15, 3719-3735.

Olson, C.L. (1974). Comparative robustness of six tests in multivariate analysis of variance. Journal of the American Statistical Association, 69, 894-908.

Robey, R. R., & Barcikowski, R. S. (1992). Type I error and the number of iterations in Monte Carlo studies of robustness. British Journal of Mathematical and Statistical Psychology, 45, 283-288.

Stevens, J. P. (1996). Applied multivariate statistics for the social sciences (3rd ed.). Mahwah, NJ: Erlbaum.

Subrahmaniam, K., & Subrahmaniam, K. (1973). On the multivariate Behrens-Fisher problem. Biometrika, 60, 107-111.

Yao, Y. (1965). An approximate degrees of freedom solution to the multivariate Behrens-Fisher problem. Biometrika, 52, 139-142.

Tables Figures, & Appendices

Table 1. Overall chi-square(χ^2 , df=540) and percent bias ($B_{\%}$) results on Type I error control of multivariate tests on means

	χ^2	Min $B_{\%}$	Max $B_{\%}$	$\overline{B}_{\%}$	$s_{\overline{B}_{\%}}$
Hotelling	2291426.84 [a]	-89	1611	96	11.5
James1	509488.00 [a]	-10	1056	57	5.2
James2	120736.97 [a]	-12	662	22	2.6
Yao	292519.99 [a]	-46	875	31	4.2
Johanson	241381.95 [a]	-13	864	33	3.7
Nel	45222.75 [a]	-98	405	-9	1.7
Kim	116189.53 [a]	-83	408	8	2.7

Note: [a] =p<.001

Table 2. Summary chi-square (χ^2 , df=90) and percent bias ($B_{\%}$) results on Type I error control of multivariate tests on means as a function of the magnitude of the correlations in $\mathbf{P}_2$.

ρ		χ^2	Min $B_{\%}$	Max $B_{\%}$	$\overline{B}_{\%}$	$s_{\overline{B}_{\%}}$
.0	Hotelling	96.72	-10	10	0	.5
	James1	37567.84 [a]	-6	375	45	8.1
	James2	4794.09 [a]	-7	142	15	3.0
	Yao	24941.08 [a]	-46	440	25	7.2
	Johanson	12962.38 [a]	-8	257	23	5.0
	Nel	3663.97 [a]	-98	8	-14	2.6
	Kim	19546.60 [a]	-53	390	12	6.7
.1	Hotelling	363.66 [a]	-9	27	3	.9
	James1	39866.49 [a]	-8	392	56	8.4
	James2	5386.21 [a]	-8	153	15	3.2
	Yao	25482.77 [a]	-45	449	25	7.3
	Johanson	14312.32 [a]	-10	269	24	5.2
	Nel	3541.31 [a]	-97	8	-14	2.5
	Kim	19297.53 [a]	-52	381	12	6.6
.3	Hotelling	11510.92 [a]	-28	142	22	4.7
	James1	45978.06 [a]	-7	455	48	9.1
	James2	6925.99 [a]	-7	194	16	3.7
	Yao	31383.76 [a]	-39	513	27	8.2
	Johanson	17573.93 [a]	-8	324	25	5.9
	Nel	3648.57 [a]	-92	12	-14	2.5
	Kim	19733.42 [a]	-50	388	-11	6.7
.5	Hotelling	89728.38 [a]	-57	401	64	12.9
	James1	62635.72 [a]	-10	584	52	10.9
	James2	11026.87 [a]	-12	272	19	4.7
	Yao	43517.76 [a]	-35	633	29	9.7
	Johanson	26424.89 [a]	-13	435	29	7.3
	Nel	3771.16 [a]	-84	20	-14	2.6
	Kim	20639.06 [a]	-46	408	10	6.9
.7	Hotelling	448624.20 [a]	-75	872	155	28.2
	James1	98281.91 [a]	-4	739	62	13.8
	James2	21492.31 [a]	-8	383	25	6.6
	Yao	61932.73 [a]	-43	729	34	11.6
	Johanson	45791.20 [a]	-9	562	37	9.7
	Nel	4912.15 [a]	-90	67	-8	3.3
	Kim	19751.95 [a]	-57	378	6	6.8
.9	Hotelling	1741102.95 [a]	-89	1611	334	53.7
	James1	225157.95 [a]	-9	1056	87	21.2
	James2	71111.50 [a]	-9	662	42	12.2
	Yao	105261.87 [a]	-12	875	43	15.1
	Johanson	124317.22 [a]	-10	864	57	16.1
	Nel	25685.57 [a]	-97	405	13	7.7
	Kim	17221.97 [a]	-83	318	-5	6.4

Note: [a] =p<.001

Table 3. Summary chi-square (χ^2 , df=180) and percent bias ($B_{\%}$) results on Type I error control of multivariate tests on means as a function of p.

p		χ^2	Min $B_{\%}$	Max $B_{\%}$	$\overline{B}_{\%}$	$s_{\overline{B}_{\%}}$
1	Hotelling	232890.75[a]	-42	770	55	11.0
	James1	64049.50[a]	-10	419	41	5.3
	James2	12419.85[a]	-12	232	15	2.5
	Yao	32101.46[a]	-46	321	19	4.1
	Johanson	18455.82[a]	-13	272	19	3.0
	Nel	6465.34[a]	-86	64	-11	1.8
	Kim	1909.13[a]	-57	240	5	3.3
2	Hotelling	756793.92[a]	-75	1326	99	19.8
	James1	163912.10[a]	-9	784	57	8.8
	James2	37092.50[a]	-9	457	22	4.4
	Yao	90050.82[a]	-38	615	31	6.9
	Johanson	72262.82[a]	-10	592	33	6.1
	Nel	13262.35[a]	-95	220	-9	2.7
	Kim	37866.03[a]	-69	329	7	4.7
3	Hotelling	1301742.16[a]	-89	1611	134	25.8
	James1	281526.41[a]	-8	1056	72	11.7
	James2	71224.62[a]	-8	662	29	6.1
	Yao	170367.71[a]	-34	875	42	9.5
	Johanson	150663.31[a]	-10	864	46	8.8
	Nel	25495.05[a]	-98	405	-5	3.9
	Kim	59226.37[a]	-83	409	11	5.9

Note: [a] =p<.001

Table 4. Summary chi-square (χ^2 , df=90) and percent bias ($B_{\%}$) results on Type I error control of multivariate tests on means as a function of n_I

n_I		χ^2	Min $B_{\%}$	Max $B_{\%}$	$\overline{B}\%$	$s_{\overline{B}\%}$
p+1	Hotelling	445543.61 [a]	-75	1611	108	30.4
	James1	469698.13 [a]	12	1056	245	30.0
	James2	115987.92 [a]	-3	662	106	12.2
	Yao	277468.24 [a]	-46	875	138	21.1
	Johanson	228867.37 [a]	2	864	154	16.6
	Nel	31973.36 [a]	-98	405	-38	7.7
	Kim	109948.22 [a]	-57	408	85	13.4
2p	Hotelling	408523.95 [a]	-84	1566	102	29.2
	James1	37298.15 [a]	4	250	43	5.3
	James2	4319.55 [a]	5	95	23	2.1
	Yao	14151.20 [a]	22	134	37	4.3
	Johanson	11862.23 [a]	-3	160	37	3.5
	Nel	10268.78 [a]	-79	202	-10	4.8
	Kim	2826.83 [a]	-83	43	-7	2.5
4p	Hotelling	375447.34 [a]	-85	1534	96	28.1
	James1	2179.83 [a]	-3	45	18	1.3
	James2	174.33 [a]	-7	15	3	.6
	Yao	639.87 [a]	-10	28	7	1.0
	Johanson	397.85 [a]	-8	25	5	.8
	Nel	2310.25 [a]	-50	76	-2	2.3
	Kim	2227.86 [a]	-67	6	-16	1.6
10p	Hotelling	359819.93 [a]	-86	1513	92	27.6
	James1	154.94 [a]	-9	15	3	.5
	James2	98.11	-12	9	0	.5
	Yao	105.94	-11	11	1	.5
	Johanson	96.77	-13	9	-1	.5
	Nel	422.69 [a]	-26	27	0	1.0
	Kim	764.15 [a]	-35	7	-9	1.0
20p	Hotelling	352263.03 [a]	-89	1519	90	27.3
	James1	72.55	-8	15	1	.4
	James2	72.91	-8	13	0	.4
	Yao	72.47	-8	13	0	.4
	Johanson	76.94	-10	12	-1	.4
	Nel	140.00 [a]	-45	18	0	.6
	Kim	294.36 [a]	-19	6	-5	.6
40p	Hotelling	349828.98 [a]	-88	1499	90	27.2
	James1	84.41	-10	9	1	.4
	James2	84.14	-10	9	0	.4
	Yao	82.28	-10	9	0	.4
	Johanson	80.49	-10	9	0	.4
	Nel	107.67	-12	13	0	.5
	Kim	128.11 [b]	-14	9	-2	.5

Note: [a] =p<.001
[b]=p<.01

Table 5. Summary chi-square (χ^2 , df=108) and percent bias ($B_{\%}$) results on Type I error control of multivariate tests on means as a function of n_1: n_2.

n_1: n_2		χ^2	Min $B_{\%}$	Max $B_{\%}$	$\bar{B}_{\%}$	$s_{\bar{B}_{\%}}$
1:1	Hotelling	9828.24[a]	-8	299	14	14.0
	James1	103486.62[a]	-8	686	66	65.7
	James2	19690.75[a]	-8	360	25	24.6
	Yao	3742.90[a]	-46	191	-1	-1.5
	Johanson	39412.60[a]	-10	510	34	34.1
	Nel	2349.02[a]	-48	88	-3	-2.6
	Kim	2790.05[a]	-57	8	-14	-14.3
1:2	Hotelling	416600.24[a]	-7	1029	159	21.2
	James1	129294.32[a]	-9	877	70	12.9
	James2	31005.39[a]	-9	504	28	6.6
	Yao	56487.60[a]	-10	677	39	8.9
	Johanson	61388.84[a]	-9	682	41	9.2
	Nel	18432.67[a]	-72	405	9	5.4
	Kim	16627.12[a]	-66	218	13	5.1
1:4	Hotelling	1852183.37[a]	-10	1611	353	43.4
	James1	251521.72[a]	-9	1056	103	17.7
	James2	67460.96[a]	-12	662	45	9.6
	Yao	226395.77[a]	-11	875	95	17.0
	Johanson	133506.68[a]	-13	864	67	13.3
	Nel	11354.13[a]	-98	226	-5	4.3
	Kim	91670.44[a]	-83	408	38	11.7
2:1	Hotelling	8459.33[a]	-89	10	-27	2.7
	James1	21681.20[a]	-10	218	33	5.0
	James2	2240.68[a]	-10	71	10	1.7
	Yao	2287.83[a]	-12	85	9	1.8
	Johanson	6096.05[a]	-10	133	15	2.8
	Nel	7717.30[a]	-93	8	-24	2.7
	Kim	4581.60[a]	-45	119	7	2.7
4:1	Hotelling	4355.66[a]	-68	10	-18	2.0
	James1	3504.13[a]	-9	96	13	2.1
	James2	339.19[a]	-9	29	3	.7
	Yao	3605.88[a]	-8	101	12	2.2
	Johanson	977.79[a]	-10	55	5	1.2
	Nel	5369.63[a]	-97	8	-19	2.3
	Kim	520.35[a]	-24	16	-4	.9

Note: [a] =p<.001

Appendix A

Observed percent bias ($B_{\%}$) on Type I error control of multivariate tests on means -- HOTELLING

n_1	n_1:n_2	P2																	
		p=4						*p*=8						*p*=12					
	ρ	.0	.1	.3	.5	.7	.9	.0	.1	.3	.5	.7	.9	.0	.1	.3	.5	.7	.9
p+1	1:1	2	-1	5	10	15	74	-5	1	7	13	46	190	-4	1	-1	19	58	299
	1:2	0	5	17	54	128	368	-5	5	25	102	268	728	-5	9	50	138	368	1029
	1:4	10	6	47	127	309	770	-4	9	87	250	576	1326	-8	23	124	367	838	1611
	2:1	6	3	-4	-8	-22	-28	-4	4	-13	-28	-42	-57	-2	-5	-21	-38	-56	-75
	4:1	-1	-4	-3	-9	-3	-8	-5	-4	-11	-22	-31	-44	-1	0	-22	-43	-57	-63
2*p*	1:1	3	-2	4	6	26	52	-4	-2	8	16	33	98	-3	-6	2	13	36	149
	1:2	3	5	30	58	141	324	-5	8	45	99	260	655	4	16	56	143	368	883
	1:4	0	5	49	120	316	736	-3	12	85	252	597	1282	-4	27	133	385	872	1566
	2:1	-1	2	-10	-14	-29	-37	4	-5	-14	-32	-49	-66	2	0	-21	-48	-67	-84
	4:1	6	0	0	2	-20	-11	9	-2	-16	-22	-41	-44	8	-2	-16	-38	-53	-63
4*p*	1:1	-2	-2	0	-2	12	25	-4	5	4	6	15	42	-1	3	-3	2	27	68
	1:2	-7	8	24	61	126	288	0	8	44	118	260	560	1	4	68	155	358	771
	1:4	2	4	48	137	327	707	-1	16	101	265	620	1227	-6	25	132	394	840	1534
	2:1	6	1	-2	-8	-24	-32	1	-6	-17	-37	-57	-70	4	-4	-25	-52	-71	-85
	4:1	-7	-2	-3	-3	-7	-14	-5	4	-14	-20	-38	-41	-1	-1	-16	-42	-57	-64
10*p*	1:1	7	-4	7	2	4	9	-7	-3	-5	-1	3	9	4	-1	3	-2	17	23
	1:2	-7	8	15	57	131	275	-2	12	45	112	237	507	3	9	62	173	359	723
	1:4	-1	3	50	133	319	697	2	19	92	265	621	1188	-2	24	136	393	858	1513
	2:1	9	-4	-7	-20	-28	-41	-8	3	-23	-35	-56	-74	10	-1	-23	-49	-75	-86
	4:1	3	-6	-3	-7	-8	-22	5	-5	-9	-23	-30	-46	0	0	-18	-36	-57	-68
20*p*	1:1	1	4	-3	5	3	0	-1	0	-4	8	10	11	-3	-8	5	3	9	14
	1:2	1	3	24	56	122	246	3	7	39	115	238	468	7	13	67	162	339	691
	1:4	-1	-3	42	148	320	668	-10	16	98	271	625	1184	-3	17	142	399	849	1519
	2:1	0	-2	-5	-17	-35	-37	6	-4	-15	-35	-58	-75	-2	-9	-22	-57	-74	-89
	4:1	10	2	4	-4	-14	-21	1	-1	-9	-22	-33	-48	-1	-4	-24	-41	-57	-62
40*p*	1:1	7	2	6	2	8	9	0	8	-7	5	8	5	-2	-2	-4	0	-2	3
	1:2	-2	-2	21	54	118	240	1	8	48	112	240	495	6	16	56	158	355	676
	1:4	4	6	52	138	320	680	5	12	98	262	606	1181	-1	23	137	401	862	1499
	2:1	2	-5	2	-29	-27	-42	-2	0	-16	-40	-63	-70	1	2	-28	-52	-71	-88
	4:1	3	-2	-1	-19	-9	-10	-3	-5	-12	-25	-32	-44	2	-1	-24	-39	-59	-67

15

Appendix A -- continued

Observed percent bias (B) on Type I error control of multivariate tests on means --JAMES1

n_1	n_1:n_2	P2																	
		p=4						p=8						p=12					
	ρ	.0	.1	.3	.5	.7	.9	.0	.1	.3	.5	.7	.9	.0	.1	.3	.5	.7	.9
p+1	1:1	171	172	178	184	194	279	231	249	250	269	310	510	291	299	299	312	395	686
	1:2	134	123	142	175	221	325	172	194	202	271	398	617	208	221	275	340	505	877
	1:4	226	213	223	270	312	419	288	304	364	439	537	784	375	392	455	584	739	1056
	2:1	133	131	111	108	84	75	176	182	145	115	103	102	215	218	190	158	118	105
	4:1	64	51	56	41	35	21	69	73	62	49	30	12	96	89	73	56	20	15
2p	1:1	54	54	54	59	76	94	60	56	64	75	87	130	59	58	71	76	97	172
	1:2	55	56	68	73	92	116	46	65	65	76	121	165	73	65	73	93	127	210
	1:4	84	85	107	109	118	139	103	101	120	127	155	207	114	118	124	165	194	250
	2:1	51	55	36	47	24	24	61	52	47	43	33	29	61	68	47	42	32	28
	4:1	28	23	25	26	15	11	27	29	17	26	7	6	33	27	25	14	14	4
4p	1:1	11	10	12	8	20	24	7	18	18	14	19	25	14	18	14	12	30	45
	1:2	11	19	24	16	25	25	16	15	18	30	29	31	14	9	25	19	31	44
	1:4	24	23	33	25	38	40	27	30	35	34	38	38	23	33	33	35	42	45
	2:1	17	7	16	21	17	12	13	14	17	12	8	8	17	16	7	6	10	3
	4:1	3	7	5	5	1	2	-1	10	-2	2	0	0	8	11	10	-2	-1	-3
10p	1:1	9	-2	8	2	3	2	-5	-2	-4	0	2	-3	6	2	5	-2	13	5
	1:2	-3	10	-5	-2	8	7	1	11	3	5	3	4	9	4	5	8	-2	7
	1:4	4	4	11	-8	5	15	6	6	5	1	8	1	4	6	8	3	1	13
	2:1	9	0	0	3	4	2	1	5	-2	1	2	-8	12	5	8	10	1	3
	4:1	0	-1	-1	5	1	-9	6	0	2	-7	2	5	0	3	8	3	-1	2
20p	1:1	1	4	-3	5	2	-2	0	0	-3	7	6	2	-2	-8	5	1	5	3
	1:2	3	3	1	-5	-2	-2	1	3	-1	1	-2	-4	5	3	0	5	-4	-3
	1:4	0	-5	-2	3	4	2	-6	3	-5	6	15	-2	-3	-3	-1	-6	-3	2
	2:1	4	-2	5	-1	1	-1	7	-1	2	7	-1	1	1	-4	3	3	-3	3
	4:1	8	1	6	-1	-3	0	-1	-2	2	1	7	0	-1	0	-6	-2	-2	2
40p	1:1	7	2	6	2	8	7	1	9	-7	5	6	2	-2	-2	-4	-1	-3	-2
	1:2	1	-3	0	-4	-1	-9	1	2	5	-1	-3	1	5	6	-2	1	9	1
	1:4	2	-3	7	4	0	1	4	-1	6	-9	-2	-3	3	0	-2	-1	3	-2
	2:1	5	-3	9	-10	2	-2	0	3	5	-5	-1	1	4	9	-2	-2	-1	2
	4:1	6	-2	3	-4	4	-1	-3	-5	-2	0	1	-6	-1	4	-1	-3	1	6

Appendix A -- continued

Observed percent bias ($B_{\%}$) on Type I error control of multivariate tests on means --JAMES2

n_1	n_1:n_2	P₂																	
		p=4						p=8						p=12					
	ρ	.0	.1	.3	.5	.7	.9	.0	.1	.3	.5	.7	.9	.0	.1	.3	.5	.7	.9
p+1	1:1	63	60	64	68	72	125	86	91	94	102	130	257	101	110	108	123	164	360
	1:2	47	42	51	71	95	171	53	65	71	113	177	337	70	78	101	132	233	504
	1:4	96	91	89	119	156	232	112	127	161	200	268	457	142	153	194	272	383	662
	2:1	47	45	33	35	21	25	52	61	45	34	28	31	71	68	70	55	31	33
	4:1	20	11	23	11	12	10	16	22	18	9	6	1	29	26	19	12	-3	4
2p	1:1	12	11	12	13	28	35	10	14	20	25	27	46	13	10	16	20	25	61
	1:2	18	17	24	24	37	48	8	18	17	17	48	57	22	21	24	24	39	72
	1:4	25	26	41	39	47	58	34	30	34	40	53	83	32	44	36	52	72	95
	2:1	13	17	0	14	-2	2	20	13	12	11	11	6	16	23	10	10	6	1
	4:1	9	4	5	9	3	7	8	8	-4	11	-5	1	8	7	6	0	5	-1
4p	1:1	-2	0	1	-4	2	7	-1	7	5	2	0	1	1	5	-3	-1	12	15
	1:2	2	7	9	1	7	0	1	0	3	11	7	1	0	-5	11	-5	1	2
	1:4	6	3	9	5	14	14	7	6	13	12	11	11	0	6	6	10	10	7
	2:1	5	-2	4	12	10	4	0	2	6	3	2	2	5	4	-4	-3	3	-5
	4:1	-4	0	1	1	-2	0	-7	5	-7	-1	-3	-2	2	5	4	-5	-3	-4
10p	1:1	7	-4	6	1	1	-3	-7	-3	-5	-2	-2	-8	4	-1	3	-4	8	-1
	1:2	-5	7	-6	-6	6	3	-2	9	0	3	-2	-1	6	0	0	3	-8	1
	1:4	-1	1	7	-12	1	9	2	4	1	-4	1	-4	0	3	3	-4	-4	5
	2:1	7	-3	-3	2	3	2	-2	2	-3	0	1	-9	8	3	6	8	-1	3
	4:1	-1	-1	-1	5	1	-9	4	-1	0	-8	2	5	0	2	8	2	-2	2
20p	1:1	1	4	-3	5	1	-3	-1	0	-4	6	5	1	-3	-8	5	1	4	3
	1:2	3	2	0	-5	-3	-3	0	2	-2	0	-4	-5	5	3	-1	4	-5	-5
	1:4	0	-7	-3	2	3	1	-7	2	-6	4	13	-4	-3	-4	-3	-8	-5	0
	2:1	3	-3	4	-2	0	-1	7	-1	1	6	-1	1	1	-5	3	3	-4	3
	4:1	8	1	6	-1	-3	0	-1	-3	2	1	7	0	-1	0	-6	-2	-2	1
40p	1:1	7	2	6	2	7	6	0	8	-7	5	6	2	-2	-2	-5	-1	-4	-2
	1:2	1	-3	0	-4	-1	-9	1	2	5	-1	-4	0	5	5	-2	1	9	0
	1:4	2	-3	7	4	0	0	4	-1	6	-9	-2	-4	3	-1	-2	-2	2	-3
	2:1	5	-3	9	-10	2	-3	0	3	5	-5	-1	1	3	9	-3	-2	-1	2
	4:1	6	-2	3	-4	4	-1	-3	-5	-2	0	1	-7	-1	4	-1	-3	1	6

Appendix A -- continued

Observed percent bias ($B_{\%}$) on Type I error control of multivariate tests on means -- YAO

n_1	n_1:n_2	P₂																	
		p=4						p=8						p=12					
	ρ	.0	.1	.3	.5	.7	.9	.0	.1	.3	.5	.7	.9	.0	.1	.3	.5	.7	.9
p+1	1:1	-46	-45	-39	-35	-43	-2	-37	-38	-31	-27	-6	91	-34	-27	-28	-14	13	191
	1:2	43	33	49	85	113	183	64	73	95	153	262	425	88	94	140	200	368	677
	1:4	255	237	248	280	300	321	339	362	408	462	530	615	440	449	513	633	729	875
	2:1	46	38	24	21	-5	-10	62	62	39	21	-8	-6	85	81	61	28	-7	-7
	4:1	61	51	57	45	30	6	78	76	68	51	27	2	101	90	78	59	26	4
2p	1:1	-19	-22	-14	-15	-3	11	-17	-14	-7	-1	13	41	-11	-14	-9	2	18	61
	1:2	31	31	38	41	53	44	23	41	41	47	81	61	45	43	50	64	81	78
	1:4	92	85	106	96	93	69	111	108	110	104	102	72	122	122	123	134	123	78
	2:1	24	29	7	16	-8	-12	40	31	22	10	2	-2	36	41	22	13	-1	-4
	4:1	27	17	23	23	12	4	22	27	15	23	3	3	27	25	25	16	12	0
4p	1:1	-10	-8	-7	-8	1	1	-6	2	1	-1	-1	4	-3	0	-6	-2	11	16
	1:2	8	14	17	6	13	-1	11	9	9	19	16	3	10	4	20	8	13	5
	1:4	20	21	27	16	19	14	25	26	28	21	18	11	20	26	22	21	14	9
	2:1	10	1	11	12	7	-1	7	9	9	6	2	0	12	11	-1	-1	1	-5
	4:1	1	4	4	4	-1	1	-2	8	-3	1	-2	-3	6	11	8	-2	0	-4
10p	1:1	7	-5	5	0	0	-2	-7	-4	-6	-2	-1	-9	3	-2	2	-4	9	-2
	1:2	-4	8	-5	-3	6	2	0	10	1	3	-2	-1	8	2	2	4	-8	0
	1:4	2	2	8	-11	2	8	5	5	1	-1	4	-2	2	5	5	-2	-2	5
	2:1	9	0	0	3	4	1	0	3	-3	-1	1	-10	11	5	7	8	-1	2
	4:1	0	-1	-1	5	0	-8	5	0	1	-7	1	4	0	2	8	3	-1	2
20p	1:1	1	3	-3	5	1	-4	-1	0	-5	6	5	0	-3	-8	4	1	5	3
	1:2	3	3	0	-5	-3	-4	0	3	-2	1	-5	-6	5	3	0	4	-4	-5
	1:4	0	-5	-1	2	4	-1	-7	2	-6	5	13	-4	-3	-3	-2	-8	-5	2
	2:1	4	-1	3	-2	0	-1	7	0	2	7	-2	2	1	-5	3	2	-4	3
	4:1	8	1	6	-1	-2	0	-1	-3	2	2	7	-1	-1	0	-6	-2	-2	1
40p	1:1	6	2	5	2	7	6	0	8	-7	5	7	1	-2	-2	-5	-1	-4	-3
	1:2	1	-2	0	-4	-1	-10	1	2	5	-2	-4	0	5	6	-2	1	9	1
	1:4	3	-3	7	3	0	1	4	-1	6	-9	-2	-4	3	0	-1	-1	2	-2
	2:1	5	-4	8	-10	2	-3	0	3	5	-5	-1	1	3	8	-2	-2	-1	2
	4:1	6	-2	3	-5	4	1	-3	-5	-2	0	1	-7	-1	4	-2	-3	1	5

Appendix A -- continued

Observed percent bias ($B_{\%}$) on Type I error control of multivariate tests on means --JOHANSON

n_1	n_1:n_2	P_2																	
		p=4						p=8						p=12					
	ρ	.0	.1	.3	.5	.7	.9	.0	.1	.3	.5	.7	.9	.0	.1	.3	.5	.7	.9
p+1	1:1	65	61	66	70	76	134	125	135	131	145	182	334	167	184	181	194	255	510
	1:2	58	54	65	89	116	196	95	112	115	173	261	439	126	137	181	223	363	682
	1:4	135	123	127	158	194	272	187	203	248	301	376	592	257	269	324	435	562	864
	2:1	59	55	43	45	28	30	93	96	79	60	47	50	129	133	115	94	61	58
	4:1	31	20	32	20	19	13	34	40	30	21	13	2	55	52	41	29	5	5
2p	1:1	13	12	13	16	33	42	22	23	29	36	44	72	26	23	35	37	48	108
	1:2	24	24	31	31	47	63	19	36	30	36	72	90	37	35	41	51	74	127
	1:4	39	42	59	56	65	78	59	53	67	71	92	127	69	77	74	105	129	160
	2:1	17	22	5	18	-1	2	32	22	18	17	14	9	33	35	23	20	12	8
	4:1	13	8	8	12	6	7	12	13	1	14	-3	0	15	12	12	4	7	-2
4p	1:1	-4	-3	-1	-5	2	9	-2	7	5	3	3	6	2	5	-2	0	16	25
	1:2	2	8	11	2	10	5	3	1	6	14	12	7	2	-3	14	3	11	12
	1:4	10	5	13	9	18	18	13	11	19	18	19	20	7	16	13	17	19	21
	2:1	6	-2	5	12	9	4	1	4	6	3	2	1	8	6	-4	-3	2	-5
	4:1	-4	0	1	1	-3	0	-8	5	-8	-2	-4	-3	2	5	4	-5	-5	-6
10p	1:1	6	-6	4	-1	0	-3	-8	-5	-7	-4	-2	-9	2	-3	1	-5	7	-2
	1:2	-6	6	-7	-6	5	3	-3	7	-1	1	-3	-2	6	0	-1	2	-9	1
	1:4	-1	0	7	-13	1	9	2	2	-1	-4	1	-5	-2	2	2	-5	-5	5
	2:1	6	-3	-3	1	2	1	-3	1	-5	-2	0	-10	7	2	5	7	-2	1
	4:1	-1	-2	-1	4	0	-10	3	-3	-1	-8	1	5	-2	1	6	2	-2	1
20p	1:1	0	3	-4	3	1	-3	-2	-2	-6	5	4	-1	-4	-10	4	-1	3	1
	1:2	2	1	-1	-6	-3	-4	0	2	-3	-1	-6	-6	4	1	-2	4	-6	-6
	1:4	-1	-7	-3	1	3	0	-8	1	-7	4	12	-5	-4	-5	-4	-9	-5	-1
	2:1	2	-3	4	-3	0	-2	6	-2	0	5	-2	0	-1	-7	3	2	-4	2
	4:1	8	0	6	-1	-3	-1	-1	-3	2	1	5	0	-1	-1	-7	-2	-2	1
40p	1:1	6	2	5	1	7	5	0	8	-8	4	5	1	-3	-3	-5	-3	-4	-3
	1:2	1	-3	-1	-4	-1	-9	0	1	4	-2	-5	0	4	5	-2	0	9	-1
	1:4	2	-4	6	3	-1	0	3	-2	5	-10	-3	-5	2	-1	-3	-3	0	-3
	2:1	5	-4	8	-10	1	-3	-1	2	4	-5	-1	0	2	7	-3	-3	-2	1
	4:1	6	-2	3	-5	4	-1	-3	-6	-3	0	1	-7	-2	4	-2	-4	1	5

Appendix A -- continued

Observed percent bias ($B_{\%}$) on Type I error control of multivariate tests on means --NEL

n_1	n_1:n_2	P_2																	
		p=4						p=8						p=12					
	ρ	.0	.1	.3	.5	.7	.9	.0	.1	.3	.5	.7	.9	.0	.1	.3	.5	.7	.9
p+1	1:1	-48	-46	-45	-44	-48	-27	-31	-34	-33	-39	-27	32	-26	-21	-31	-30	-20	81
	1:2	-57	-62	-52	-42	-19	64	-71	-70	-61	-37	24	220	-72	-67	-54	-30	64	405
	1:4	-86	-81	-82	-71	-48	5	-94	-95	-92	-83	-51	105	-98	-97	-92	-82	-30	226
	2:1	-58	-58	-62	-57	-62	-56	-70	-67	-73	-74	-80	-83	-70	-70	-76	-84	-90	-93
	4:1	-26	-32	-25	-45	-54	-66	-32	-29	-49	-64	-79	-92	-30	-42	-55	-73	-90	-97
2p	1:1	-16	-17	-14	-17	-11	6	-13	-12	-5	-7	4	48	-8	-11	-9	-6	5	88
	1:2	-23	-20	-15	-2	19	60	-28	-18	-13	0	59	129	-19	-19	-3	19	67	202
	1:4	-45	-44	-32	-21	-1	32	-35	-37	-28	-7	22	95	-43	-33	-25	8	56	139
	2:1	-27	-22	-35	-32	-46	-51	-20	-26	-39	-52	-58	-70	-22	-19	-39	-60	-69	-79
	4:1	-4	-13	-18	-20	-39	-52	-7	-8	-27	-34	-53	-71	-4	-11	-22	-44	-56	-75
4p	1:1	-7	-7	-5	-10	0	13	-6	4	1	-3	2	21	-2	0	-6	-4	16	46
	1:2	-6	0	3	2	17	28	-7	-6	2	20	31	54	-7	-10	12	13	40	76
	1:4	-6	-8	0	0	12	24	-4	-7	4	12	23	36	-15	-3	-1	14	28	51
	2:1	-4	-11	-7	-7	-17	-30	-7	-11	-14	-20	-36	-36	-3	-8	-23	-33	-43	-50
	4:1	-8	-4	-3	-11	-20	-32	-9	2	-15	-18	-29	-45	-1	1	-6	-25	-38	-49
10p	1:1	7	-5	5	0	2	4	-7	-3	-6	-2	1	3	3	-2	2	-4	13	15
	1:2	-6	6	-6	-3	12	13	-4	8	1	7	7	18	6	0	4	13	9	27
	1:4	-2	-1	6	-12	5	17	1	2	1	0	10	8	-3	1	4	2	6	25
	2:1	6	-3	-6	-6	-7	-12	-2	0	-9	-10	-18	-26	7	2	1	-5	-18	-22
	4:1	-1	-2	-2	1	-10	-23	4	-2	-3	-13	-10	-12	-1	1	4	-5	-14	-20
20p	1:1	1	4	-4	5	2	-1	-1	0	-4	7	7	7	-3	-8	4	1	7	9
	1:2	3	2	1	-4	1	3	0	2	-1	3	1	3	5	3	1	8	4	8
	1:4	0	-7	-3	3	5	3	-7	2	-6	6	18	3	-4	-4	-2	-4	-1	9
	2:1	3	-3	3	-6	-7	-8	6	-2	-2	2	-10	-10	1	-6	1	-3	-12	-7
	4:1	8	0	6	-3	-7	-7	-1	-3	1	-2	-1	-9	-1	0	-8	-5	-10	-10
40p	1:1	7	2	6	2	8	8	0	8	-7	5	7	4	-2	-2	-5	-1	-3	2
	1:2	1	-3	0	-4	1	-6	1	2	6	-1	0	7	5	6	-2	3	13	6
	1:4	2	-3	7	4	2	3	4	-1	6	-9	0	-2	3	-1	-1	0	5	3
	2:1	5	-4	8	-12	-2	-5	0	3	4	-7	-4	-5	3	8	-4	-5	-6	-3
	4:1	6	-2	3	-6	3	-6	-3	-6	-3	-1	-1	-12	-1	4	-2	-4	-2	0

Appendix A -- continued

Observed percent bias ($B_{\%}$) on Type I error control of multivariate tests on means --KIM

n_1	n_1:n_2	P2																	
		p=4						p=8						p=12					
	ρ	.0	.1	.3	.5	.7	.9	.0	.1	.3	.5	.7	.9	.0	.1	.3	.5	.7	.9
p+1	1:1	-53	-52	-50	-46	-57	-44	-41	-41	-40	-42	-38	-42	-37	-31	-34	-33	-36	-39
	1:2	58	56	60	87	91	95	93	100	109	128	157	143	122	118	143	154	198	218
	1:4	240	221	220	224	211	161	303	321	329	321	302	235	390	381	388	408	378	318
	2:1	60	62	42	31	5	-26	93	93	65	47	6	-40	119	118	101	66	8	-45
	4:1	9	3	16	5	15	8	-2	8	2	10	16	5	6	5	11	16	12	13
2p	1:1	-26	-24	-19	-24	-21	-33	-21	-15	-13	-16	-18	-50	-11	-16	-12	-17	-28	-57
	1:2	19	15	21	15	6	-23	5	18	8	-9	-8	-56	20	18	8	-6	-29	-66
	1:4	35	32	43	23	4	-32	25	20	11	-4	-33	-69	16	20	-5	-19	-43	-83
	2:1	10	20	1	11	-7	-20	16	13	13	8	5	-20	13	22	11	14	1	-32
	4:1	-9	-15	-15	-4	-5	-1	-17	-22	-22	-2	-6	5	-20	-24	-19	-11	1	0
4p	1:1	-8	-8	-7	-13	-13	-26	-7	3	-4	-9	-16	-39	-3	-1	-9	-12	-16	-39
	1:2	-3	1	0	-12	-14	-36	-8	-8	-12	-11	-26	-54	-7	-16	-9	-30	-40	-58
	1:4	-9	-13	-15	-22	-28	-41	-19	-20	-22	-31	-44	-56	-32	-25	-32	-37	-45	-67
	2:1	-1	-8	-1	6	2	-9	-6	-10	-2	0	-1	-13	-5	-5	-8	-4	4	-18
	4:1	-18	-12	-12	-7	-8	-2	-23	-17	-19	-10	-4	-2	-18	-17	-11	-16	-6	1
10p	1:1	7	-4	5	0	-6	-14	-7	-5	-7	-6	-10	-25	4	-2	0	-9	-4	-22
	1:2	-9	5	-12	-10	-8	-13	-8	3	-6	-11	-20	-26	1	-5	-7	-10	-28	-26
	1:4	-14	-13	-9	-24	-21	-12	-13	-12	-17	-25	-20	-35	-19	-15	-19	-28	-31	-29
	2:1	1	-6	-5	-5	2	-5	-9	0	-6	-2	2	-10	2	-1	1	6	1	-6
	4:1	-6	-7	-8	-3	-5	-14	-6	-12	-6	-6	3	7	-9	-8	-3	-1	-3	1
20p	1:1	1	2	-4	3	0	-6	0	-1	-4	4	2	-11	-2	-8	4	-1	-1	-11
	1:2	-1	-2	-3	-7	-11	-16	-1	-3	-7	-6	-15	-17	3	-2	-6	-1	-15	-19
	1:4	-8	-15	-11	-8	-5	-10	-17	-7	-14	-6	-7	-13	-13	-14	-17	-16	-16	-15
	2:1	-2	-6	1	-5	-2	-3	2	-4	-2	6	-2	-2	-4	-9	1	0	-5	1
	4:1	5	-5	5	-1	-3	-1	-5	-6	2	-4	3	-1	-6	-6	-11	-7	-3	3
40p	1:1	7	2	6	-1	5	1	0	8	-7	4	2	-6	-3	-2	-5	-1	-6	-9
	1:2	-2	-5	-4	-8	-8	-9	-3	0	1	-4	-7	-7	3	3	-5	-4	2	-4
	1:4	-1	-3	0	0	-3	-6	0	-6	-2	-13	-9	-7	-5	-6	-13	-7	-8	-6
	2:1	2	-5	9	-14	1	-2	-2	1	2	-6	0	-3	1	6	-4	-2	0	0
	4:1	3	-3	0	-10	-1	2	-6	-7	-5	-2	0	-5	-4	-2	-5	-3	-1	7

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC

TM 028438

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Type I error control of two-group multivariate tests on means under conditions of heterogeneous correlation structure

Author(s): Rachel T. Fouladi

Corporate Source: Dept of Educational Psychology University of Texas at Austin

Publication Date: April 1998

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
Sample	Sample	Sample
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1	2A	2B
Level 1	Level 2A	Level 2B
☑	☐	☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here,→ please

Signature: Rachel Tanya Fouladi

Printed Name/Position/Title: Rachel Tanya Fouladi / Asst Prof

Organization/Address: SZB 504, UT Austin, Austin TX 78712-1296

Telephone: 512 471-4155 FAX: 512 471-1288

E-Mail Address: rachel.fouladi@mail.utexas.edu Date: 4/15/98

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

THE UNIVERSITY OF MARYLAND
ERIC CLEARINGHOUSE ON ASSESSMENT AND EVALUATION
1129 SHRIVER LAB, CAMPUS DRIVE
COLLEGE PARK, MD 20742-5701
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2nd Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com